

Primary Creep in Sn3.8Ag0.7Cu Solder. Part I: Theory, Experiments, and Data Reduction

D.R. SHIRLEY² and J.K. SPELT^{1,2,3}

1.—Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Rd., Toronto, ON, M5S 3G8, Canada. 2.—Department of Materials Science and Engineering, University of Toronto, 184 College Street, Toronto, ON, M5S 3E4, Canada. 3.—e-mail: spelt@mie.utoronto.ca

This article presents constant-load creep and stress relaxation data for Sn3.8Ag0.7Cu spanning a range of strain rates $10^{-8} \text{ s}^{-1} < \dot{\epsilon} < 10^{-4} \text{ s}^{-1}$, and temperatures 25°C, 75°C, and 100°C. Creep and stress relaxation measurements showed that transient creep caused faster strain rates during stress relaxation for a given stress than the corresponding minimum creep rate from constant-load creep tests. The extent of strain hardening during primary creep was a function of temperature and strain rate. Data reduction incorporated a fast Fourier transform method to remove spurious data from stress relaxation corresponding to the period of partial strain relaxation during loading.

Key words: Tin–silver–copper (Sn–Ag–Cu) solder, transient creep, stress relaxation, reliability

INTRODUCTION

The creep of SnAgCu solder alloy is an important factor governing the reliability of solder joints. The solidus temperature (T_s) of Sn3.8Ag0.7Cu is 217°C (490 K), so the homologous temperature ($T_h = T/T_s$) defined in absolute temperature terms is relatively high ($T_h > 0.6$) even at room temperature. Consequently, solder joints at typical service temperatures and during thermal cycling tests are subject to significant creep deformation, even at low stresses. Creep is time-dependent plastic deformation that may readily occur due to thermal or applied stresses below the proof stress of a metal. Creep strain is a function of temperature, strain rate, and stress. Historically, creep in SnAgCu has been studied at relatively high strain rates and stresses that are often not representative of conditions in actual solder joints.

The total strain exhibited during creep testing can be partitioned into elastic and inelastic deformation (Eq. 1). The inelastic deformation is comprised of time-independent (plastic) and time-dependent (creep) deformation. The fluctuating

stresses in typical solder joints usually do not exceed yield,^{1,2} hence the plastic or inelastic strain will refer solely to creep in the following analyses.

$$\epsilon_{\text{total}} = \epsilon_{\text{elastic}} + \epsilon_{\text{inelastic}} \quad (1)$$

Creep strain contributions in metals under constant load and temperature are partitioned into three stages: (1) primary (transient) creep, with a decelerating strain rate as the metal strain hardens; (2) secondary (steady-state) creep, where the strain hardening and dynamic recovery (softening) processes are in equilibrium and the metal deforms approximately linearly with time at an averaged minimum creep rate; and (3) tertiary (transient) creep, where voids form and the strain rate accelerates to final rupture. Tertiary creep is usually omitted in solder joint deformation modeling because it represents only deformation contributions at the onset of imminent creep rupture. Although primary and secondary creep regimes are often considered as distinct material responses to applied stress and temperature, both are strongly interconnected at high homologous temperatures.³ Given that true strains rates are additive, the total true strain rate is given by Eq. 2, where $\dot{\epsilon}_{\text{elastic}}$ is the true elastic strain rate, $\dot{\epsilon}_{\text{tr}}$ is the true transient creep strain rate and $\dot{\epsilon}_{\text{min}}$ is the true minimum

(Received December 8, 2008; accepted July 13, 2009; published online September 16, 2009)

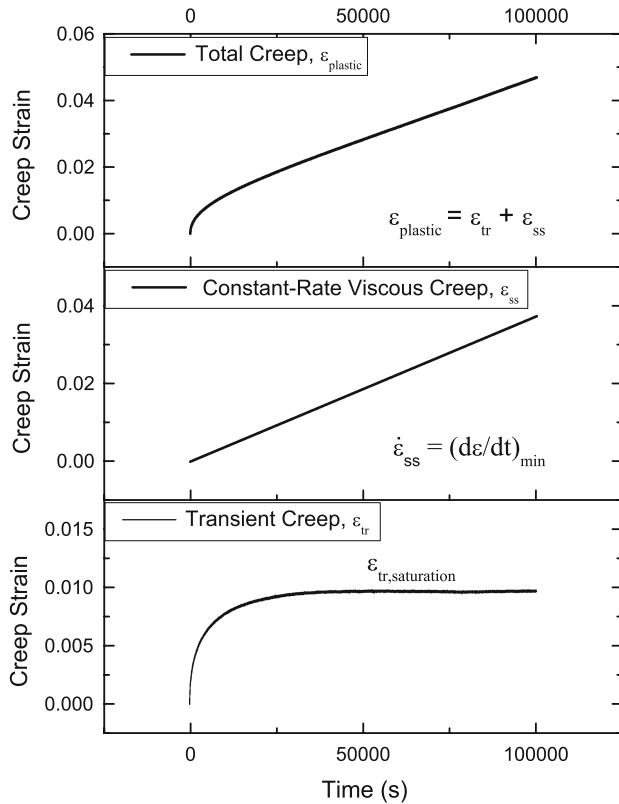


Fig. 1. Creep strain versus time. Superposition of creep components that determine the creep curve ($T = 25^\circ\text{C}$, $\sigma = 15\text{ MPa}$).

(steady-state) creep strain rate. Total creep strain is a superposition of transient and constant-rate steady state creep (ϵ_{ss}) components. Although both contribute to the creep deformation shown in Fig. 1, transient creep (ϵ_{tr}) is dominant during the primary creep stage, and steady state creep is dominant during the secondary creep stage.⁴

$$\dot{\epsilon}_{\text{total}} = \dot{\epsilon}_{\text{elastic}} + \dot{\epsilon}_{\text{tr}} + \dot{\epsilon}_{\text{min}} \quad (2)$$

During isothermal creep testing at a given stress level, the contributions of transient creep gradually saturate during primary creep, and subsequent creep contributions are dominated by steady-state creep (secondary creep). Whenever a stress change occurs within a metal, some transient creep is produced prior to the onset of steady-state creep.⁵ Stress fluctuations and, hence, transient creep are present in thermally cycled solder joints due to mismatch in the thermal expansion coefficients of adjacent materials. The isothermal dwell period of a thermal cycle is analogous to a constant displacement loading where stress relaxes over time. Although transient (primary) creep is often ignored in solder joint modeling and within the electronics packaging literature,^{6–8} it has been shown to make a significant contribution to total deformation during thermal cycling of SnAgCu solder joints.^{2,9}

This paper describes primary and secondary creep measurements made with Sn3.8Ag0.7Cu



Fig. 2. Tensile specimen (ruler units millimeters).

solder alloy under both constant-load and constant-displacement (stress relaxation) conditions at three temperatures. Stress and strain rates were selected to extend down to ranges representative of actual solder joints in service ($\dot{\epsilon} < 10^{-4}\text{ s}^{-1}$ and $\sigma < 20\text{ MPa}$). Several non-standard data reduction techniques were used to interpret the results and characterize both the primary and secondary creep regimes. The data are used in a companion paper¹⁰ to develop a creep constitutive model incorporating both transient and steady-state creep strain.

EXPERIMENTAL PROCEDURES

Bulk Tensile Specimen

The alloy investigated was the near-eutectic ternary alloy Sn3.8Ag0.7Cu: Ag 3.73, Cu 0.68, Pb 0.05, Sb 0.03, Bi 0.01, and balance Sn (in weight-percent). The bulk solder tensile specimen (Fig. 2) was designed according to international standard guidelines^{11,12} to achieve a uniform uni-axial stress in the reduced radius gauge section with a 40 mm gauge length and 8 mm diameter. Bulk bar stock of Kester Sn3.8Ag0.7Cu was melted, cast into 14 mm diameter copper pipe molds, and air cooled at $2\text{--}3^\circ\text{C/s}$, representative of cooling during reflow processing in microelectronic assembly. Each casting was scanned by non-destructive C-mode scanning acoustic microscopy (C-SAM) to inspect for voids, cracks, and foreign material inclusions.¹³ The gauge section was machined with a CNC lathe to give a 6 mm shoulder fillet radius that transitioned to a 12.7 mm diameter straight cylindrical cross-section along the balance of the tensile specimen. The ends were machined flat and soldered to copper rods (60 mm long and 12.7 mm diameter). The rods were threaded along half the length and subsequently inserted into steel sockets that connected the specimen to the load train of the constant-load creep testing frame. This prevented thread failure that otherwise occurred at elevated temperatures, when threads made directly in bulk solder deformed excessively.

Once machined, each specimen was aged at $\frac{2}{3} * T_{\text{solidus}}$ (54°C) for 16 h to relax residual stresses and provide a common starting microstructure.¹⁴ Figure 3a and b shows that the aged and untested specimen had large Sn-rich regions containing ternary eutectic Ag3Sn and Cu6Sn5 rod-like particles (some appear as small dark dots) that define the edge of β -Sn grain boundaries.

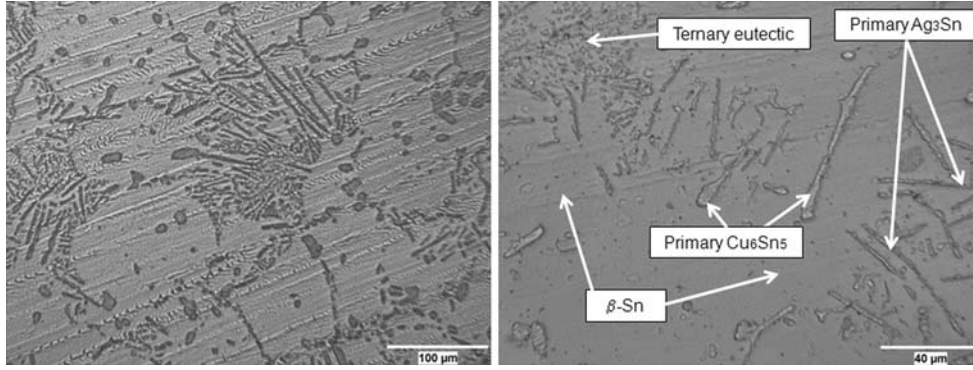


Fig. 3. Optical micrographs of multiple locations using quarter-wave plate of aged as-cast tensile specimen microstructure (no deformation). (a) Low magnification, (b) high magnification.

Experimental Setup

The isothermal constant-load creep test matrix (Table I) included tests at several stress levels at each of 25°C, 75°C, and 100°C. The constant-displacement test matrix included duplicate isothermal tests each run at 25°C, 75°C, and 100°C. The specimens in all tests were placed within an aluminum sheet enclosure, through which an opening was made to accommodate a pair of T-type thermocouples and an extensometer. Thermocouples were attached to the specimen surface near the top and bottom portions of the reduced diameter gauge section. The uniform strain within this section was measured with a clip gauge extensometer (10 mm gauge length) affixed with springs. Heating was provided by a set of rope heaters attached directly to both steel sockets that connected with the threaded copper ends of the tensile specimen. The heaters were controlled by a proportional–integral–derivative (PID) algorithm with a set of solid state relays using a temperature feedback loop to keep the temperature within $\pm 1^\circ\text{C}$ of the set point.

The data acquisition program recorded temperature, strain, and force at a sampling rate of 60 per minute and 600 per min, for creep and stress relaxation tests, respectively. Electrical noise was filtered with a low-pass finite impulse response filter with a cutoff frequency equal to one-tenth the sampling frequency.

Most of the constant-load creep tests were conducted with the dead-weight load frames by attachment of calibrated weights to the load train and the gradually lowering of a supporting platform

from below, until the weights were freely suspended from the specimen. Some additional creep tests were made with the screw-driven actuator load frame in load control mode ($\pm 2\text{ N}$). This load frame was also used in stroke-control mode for all constant-displacement stress relaxation tests.

Data Reduction

Constant Load-Creep

The start of the creep test (the ‘zero moment’) was the moment when the test load was fully applied, corresponding to the maximum measured strain rate.¹¹ During the finite loading time required to strain the specimen the stress is described by the Boltzmann integral.¹⁵ The zero moment was readily determined by our fitting the loading portion of the strain (ϵ_{total}) versus time (t) curve to a sigmoidal Boltzmann function (Eq. 3) and solving for the maximum slope, where A_1 , A_2 , t_0 , and dt are fitting parameters.

$$\epsilon_{\text{total}} = \frac{A_1 - A_2}{1 + \exp\left(\frac{t-t_0}{dt}\right)} + A_2 \quad (3)$$

Throughout the creep test, particularly at low strain rates, the extensometer measurement was subject to residual signal noise. A smooth curve was obtained by an adjacent averaging approach for particularly noisy strain-time data. Following the zero moment, the creep strain rate decelerated to an approximately linear strain rate, denoting the transition between primary and secondary creep regimes. This transition occurs gradually and is difficult to determine precisely by inspection for many creep curves; however, the definition of the saturation transient creep strain and minimum creep rate relies on the accurate determination of the transition time t_{sat} , the elapsed time to transient strain saturation. An algorithm was created to define this transition point over time, based on the partial memory sliding-window method that used a statistical hypothesis test to detect change.^{16,17} The initial reference time window was taken from the

Table I. Test Matrix for Constant Load Creep

Temperature (°C)	Engineering Tensile Stress (MPa)
25	6.5, 10.0, 15.0 ($\times 2$), 20.0
75	2.6, 7.8 ($\times 2$), 11.3, 14.8
100	2.6, 3.9, 5.2 ($\times 2$), 8.6, 12.0

end of the creep curve, in the linear portion of the secondary creep regime with subsequent windows sliding progressively backwards in time. The width of this window was selected as described below. The curves were truncated if the full data set extended into the tertiary creep regime. In such cases a similar technique was used to determine the transition from secondary to tertiary, with the initial reference window within the linear region and subsequent windows sliding forwards in time. From the central limit theorem an appropriate sample size for the statistical test was obtained from a time window divided into 30 sub-windows, where the slope of each was determined by linear regression.¹⁶ The mean of these first 30 slopes became the reference window slope. The algorithm then advanced to the next time window closer to the start of the creep curve ($t = 0$ s) by an increment of two sub-windows, adding two new recent slopes and removing the oldest two slopes from the current window. The slope of each of the 30 examples within the current window was taken by linear regression, and the mean was compared with the mean of the reference window by a two-tailed t -test with an alpha level of 0.05. If the mean of the slopes were statistically equal, two new sub-windows were added and two removed from the end of the window, and the process was repeated. When the mean slope of the current window was not statistically equal to that of the reference window, the transition time, t_{sat} , was considered to be the middle of this window. Since the gradual nature of change in the primary and secondary creep regime transition was similar between specimens, a fixed size time window was implemented. In cases where the data were noisy, a larger window size was useful for smoothing, although a window either too large or too small resulted in decreased prediction accuracy. Since the calculated transition time was sensitive to the window size, an iterative approach was used to find the best estimate. The initial transition time solution was calculated with a window size of one-half of the total elapsed time (t_{elapsed}) of the measured creep strain versus time data set. Subsequent solutions were re-run with progressively smaller windows, down to one-eighth the total elapsed time, where the iterative window size was defined by t_{elapsed}/n , where $n = 2, 3, \dots, 7, 8$. The estimate of transition time (t_{sat}) was then taken as the median of the array of all calculated times computed from each of the several window sizes.

The linear portion of the creep curve beyond the transition time was fit using linear regression to solve for the slope and y -axis intercept that corresponded to the minimum creep rate and saturated transient creep strain, respectively.

Constant Displacement-Stress Relaxation

The stress relaxation tests were loaded by movement of the screw-driven actuator crosshead at a

constant speed to a fixed displacement. However, since loading requires a finite time, during and immediately after the loading period some relaxation occurred, mostly of high frequency components (considering the strain-time response as a Fourier series).¹⁵ Consequently, pure relaxation under constant displacement starts from values less than the maximum stress at a time beyond the end of the loading period. If the tensile specimen has a relaxation time that is shorter than the loading time, then the amplitude associated with loading will dissipate faster than it is applied, and it will have decreased considerably from its true value by the time loading is complete.

The main limitation of stress relaxation is the collection of high frequency data components. The limiting factor is not the sampling data rate but the finite time taken to strain the specimen. Accordingly, this finite loading time was minimized with relatively fast constant crosshead speeds (2 cm/min at 25°C; 4 cm/min at 75°C; and 6 cm/min at 100°C). The simultaneous relaxation and application of stress during the initial loading period means that all frequency components will be at least partially present at the end of loading, regardless of crosshead speed.¹⁵

The identification of the start of pure relaxation, somewhat later than the final crosshead movement, was adapted from Ref. 15 using a fast Fourier transform (FFT) analysis and an FFT low-pass filter. The FFT analysis approximately transforms a periodic time-domain waveform, such as force versus time response during stress relaxation, into the frequency domain, as a finite summation of sinusoids of different frequencies and magnitudes.¹⁸ Although the stress relaxation is a continuous analog function of time, the data were acquired at a fixed sampling rate, making the force measurement a discrete function of time and thereby introducing a distortion termed aliasing, most acute at the highest frequencies. Aliasing was mitigated by our setting the measurement rate well above the Nyquist sampling rate.¹⁸ Similarly, the truncation of an infinite series to a finite representation of the Fourier series introduced a distortion termed leakage that affected the frequency resolution of the FFT. The leakage of the FFT was significantly reduced by use of the Hanning window function in the signal processing tools.¹⁹

The amplitude-frequency single-sided power spectrum (Fig. 4) was used to determine the start of pure relaxation using the fact that the high frequency components have relaxation periods much shorter than the loading period. Consequently, at the end of loading, the lower frequency components will have higher magnitudes and tend toward their true relaxation values, whereas the higher frequency components, due to the finite time required for crosshead movement, would have decreased considerably.¹⁵ For example, Fig. 4 shows the FFT frequency response for stress

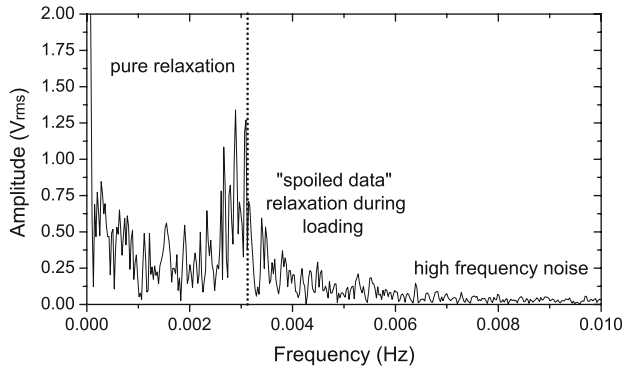


Fig. 4. Stress relaxation at 100°C, FFT frequency-domain response, $f_c = 0.0031$ Hz.

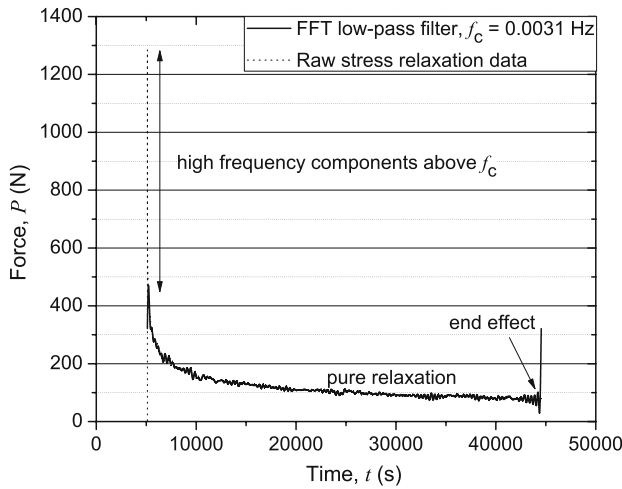


Fig. 5. Stress relaxation at 100°C, raw and filtered time-domain response.

relaxation at 100°C, where the pure signal and incomplete high frequency regimes are indicated. The point at which the amplitude decreased sharply from its maximum value, beyond the initial lobe, was used to determine the cutoff frequency at the start of pure relaxation. Once the cutoff frequency (f_c) had been identified, an FFT low-pass filter was applied to the load cell output to yield a stress versus time response without the higher frequency components. Subtraction of the highest frequency components can also effectively mitigate a large part of the signal degradation due to aliasing.¹⁵ Figure 5 shows the original and filtered time-domain representation of this stress relaxation response, where the high frequency components have been removed. The approximation of the finite Fourier series is unable to represent exactly the abrupt end of the relaxation curve resulting in the second-order 'end effect' of the filtered load-time response.²⁰ This oscillatory end effect was an artifact of the FFT approach and was omitted from the pure relaxation response.

Machine Compliance

The interpretation of the constant-displacement stress relaxation measurements required consideration of the compliance of the tensile specimen and the screw-driven actuator test machine.^{21,22} The cross-head displacement was the sum of the elastic and plastic elongation of both the tensile specimen and machine elements, including the crosshead, linkages, and load cell. If we assume that all elements of the machine are linearly elastic and that all inelastic deformation is isolated to the tensile specimen gauge length, Eq. 4 defines the total crosshead displacement over time (t).

$$\int_0^t \dot{X} dt = \frac{\sigma L_0}{E} + \frac{\sigma A_0}{K} + \varepsilon_P L_0 \quad (4)$$

where \dot{X} is crosshead speed, t is time, σ is gauge section stress, K is machine module, E is Young's modulus of solder, A_0 is initial cross-sectional area, ε_P is true plastic strain, and L_0 is initial gauge length.

In a tensile specimen with a large shoulder fillet diameter some inelastic deformation may, in fact, occur in the shoulder transition beyond the parallel gauge section. This may be incorporated as an increase in the initial gauge length. This effective gauge length was calculated with Eq. 5 by measurement of the uniform total true strain ($\varepsilon_{\text{total}}$) and the final shoulder-shoulder separation ($L_{f,\text{shoulder}}$) with dial calipers.

$$L_0 = \frac{L_{f,\text{shoulder}}}{\varepsilon_{\text{total}}} \quad (5)$$

In the case of a constant-strain (strain control) test the total true strain rate ($\dot{\varepsilon}_{\text{total}} = 0$) is zero, resulting in effectively infinite machine stiffness. However, in the case of constant displacement (stroke control) the crosshead is fixed ($\dot{X} = 0$), thus machine stiffness (K) does contribute to true plastic strain rate and the interpretation of stress relaxation results. The coupling relationship between the tensile specimen and machine is given by our taking the derivative of Eq. 4 and dividing by the gauge length. Simplification of the resultant coupling expression further by substitution of the machine modulus (M), where $M = \frac{KL_0}{A_0}$, gives Eq. 6. Thus, the combined specimen-machine elastic modulus (C) is given by Eq. 7, where the coupled elastic response acts as two springs in series.

$$\frac{\dot{X}}{L_0} = \dot{\sigma} \left(\frac{1}{E} + \frac{1}{M} \right) + \dot{\varepsilon}_P \quad (6)$$

$$\frac{1}{C} = \frac{1}{E} + \frac{1}{M} \quad (7)$$

Equation 6 is exact for load-independent machine stiffness which is a good approximation over the

stress relaxation test.²² The machine stiffness is known to vary between tests, even for identical test conditions, hence we must calculate K for each mechanical test performed, rather than assume a universal value for the machine.²¹

The machine stiffness was determined by the total deformation method. During initial loading the crosshead speed (\dot{X}) was constant, and the linear portion of the loading rate (\dot{P}) and the corresponding total strain rate ($\dot{\epsilon}_{\text{total}}$) were each calculated by linear regression. Machine stiffness was calculated from Eq. 8, which was derived from Eq. 6 for the elastic portion of initial loading with constant crosshead speed.

$$K = \left[\left(\frac{L_0}{\dot{P}} \right) * \left(\frac{\dot{X}}{L_0} - (\dot{\epsilon}_{\text{elastic}} + \dot{\epsilon}_{\text{plastic}}) \right) \right]^{-1} \quad (8)$$

RESULTS AND DISCUSSION

Constant Load Creep

The constant load creep strain versus time data for 25°C, 75°C, and 100°C are given in Figs. 6, 7, and 8, respectively (omitting the duplicate tests indicated in Table I for clarity). The figures show the transitions from primary creep to secondary creep regimes. The transient saturation creep strain, $\epsilon_{\text{tr,sat}}$, minimum creep rate, $(\frac{d\epsilon}{dt})_{\text{min}}$, and saturation transition time, t_{sat} , calculated as explained above, are shown in each figure. The total creep

strain in these curves is a superposition of two creep processes, transient, and the constant-rate steady state creep components.⁴ Although transient and steady state creep are sometimes considered to be distinct material responses to applied stress and temperature, they occur concurrently and are strongly interconnected at high homologous temperatures.³ Transient creep was the dominant component of total creep strain at the beginning of the constant load creep tests and increased gradually until it reached a maximum value ($\epsilon_{\text{tr,sat}}$) at the onset of saturation (t_{sat}). Beyond t_{sat} additional increases in total creep strain were due solely to the constant-rate creep component that gradually dominates total creep strain over time.

Consider a set of creep results for similar minimum strain rates ranging from $3 \times 10^{-6} \text{ s}^{-1}$ to $5 \times 10^{-6} \text{ s}^{-1}$ measured at different temperatures and applied stresses: Fig. 6 (25°C, $\sigma = 20 \text{ MPa}$); Fig. 7 (75°C, $\sigma = 11.3 \text{ MPa}$); and Fig. 8 (100°C, $\sigma = 8.6 \text{ MPa}$). In Fig. 6 there is a large initial slope at 25°C that corresponds to strain (also termed work), hardening during primary creep. In Figs. 7 and 8 for 75°C and 100°C, respectively, the initial slope for the same $3 \times 10^{-6} \text{ s}^{-1}$ to $5 \times 10^{-6} \text{ s}^{-1}$ strain rate range is progressively smaller, as is the absolute value of creep strain measured. This is the consequence of decreased strain hardening at higher temperatures as a result of increased thermally activated recovery processes.²³ This trend is consistent with Fig. 9 for saturated transient creep strain versus minimum creep rate. Strain hardening

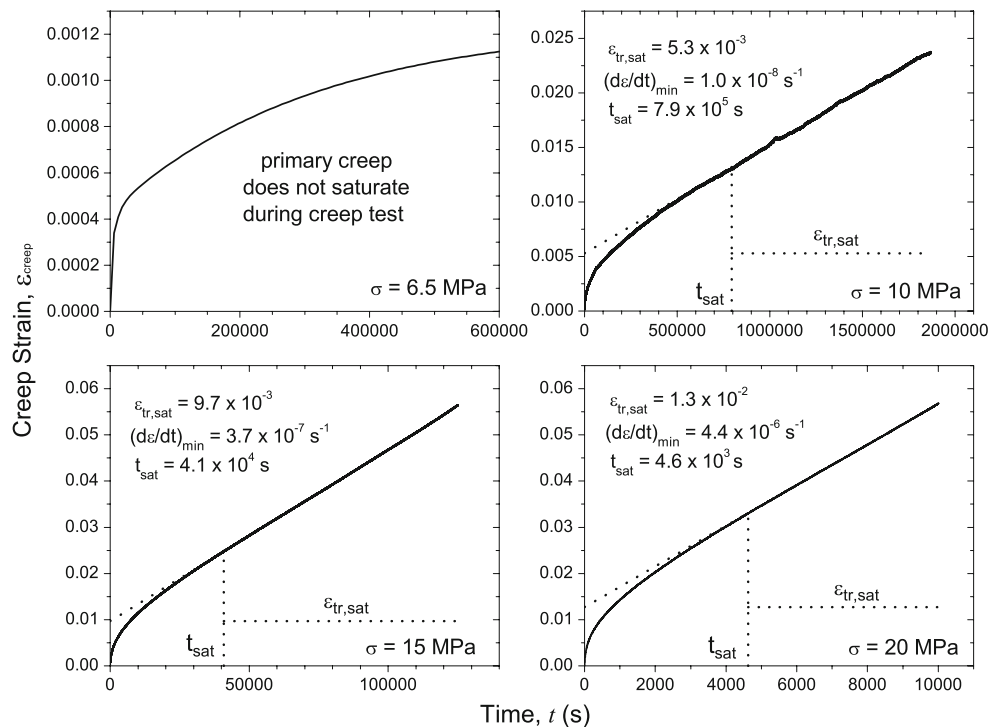


Fig. 6. Nonlinear fit to measured creep strain versus time data at 25°C.

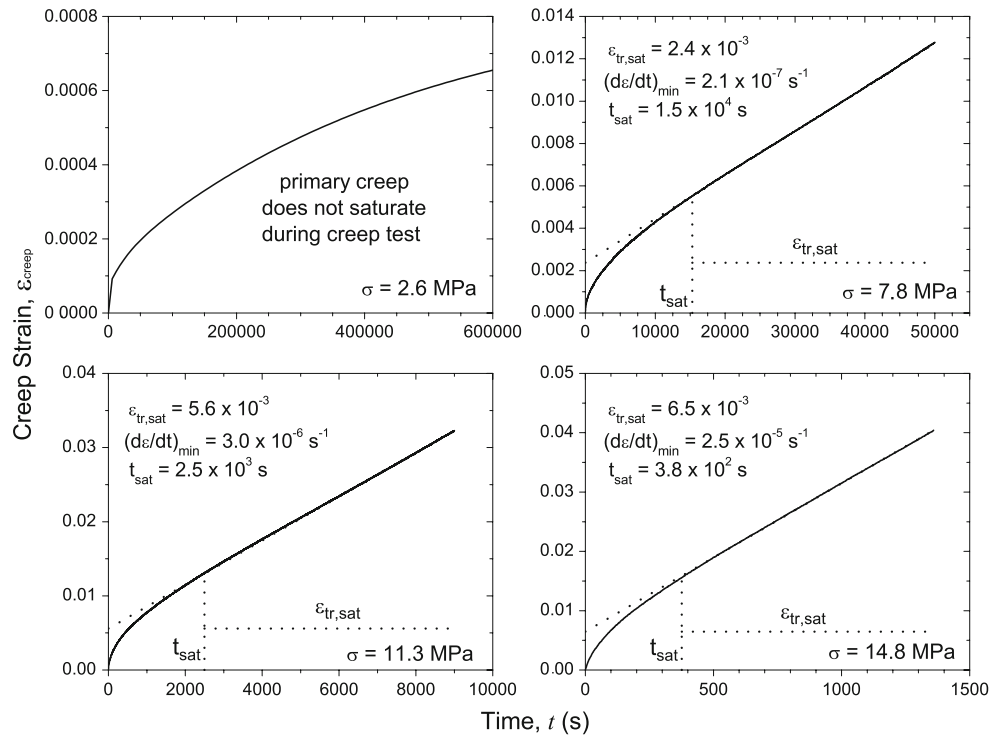


Fig. 7. Nonlinear fit to measured creep strain versus time data at 75°C.

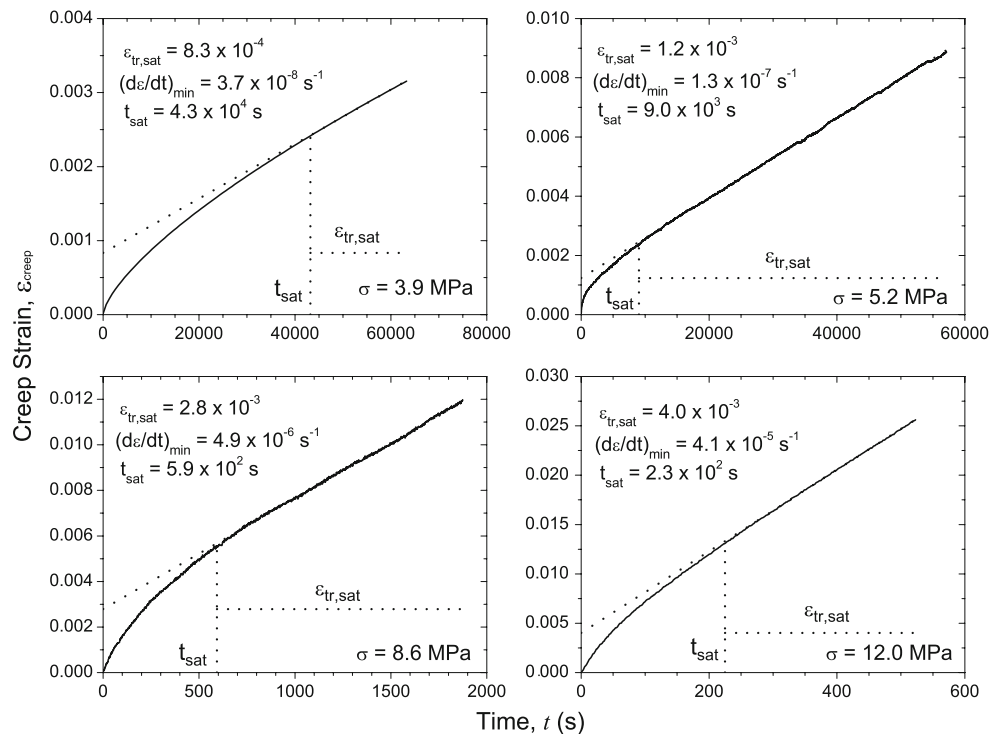


Fig. 8. Nonlinear fit to measured creep strain versus time data at 100°C.

led to larger absolute values of saturation transient creep strain at 25°C and progressively smaller amounts at 75°C and 100°C. In Fig. 10 tensile stress is plotted versus minimum creep rate.

Note that the repeated creep results in Figs. 9 and 10 for each test temperature (Table I) give similar results, indicating experimental repeatability.

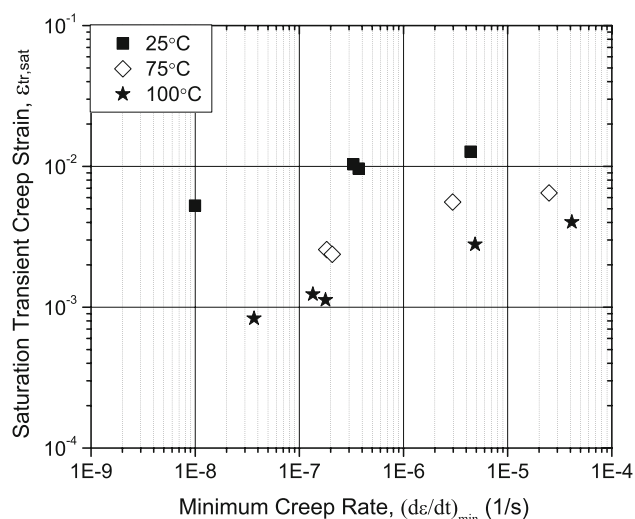


Fig. 9. Saturation transient creep strain versus minimum creep rate (25°C, 75°C, 100°C).

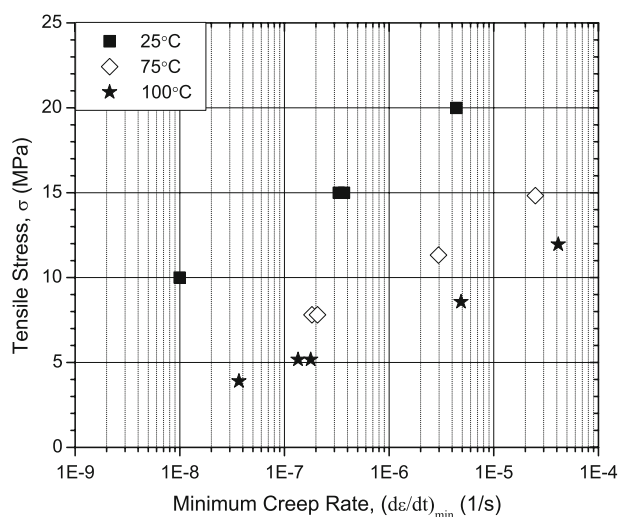


Fig. 10. Tensile stress versus minimum creep rate (25°C, 75°C, 100°C).

As widely reported in the literature, the minimum constant creep rate is a function of applied stress and temperature (Fig. 10), where faster minimum creep rates are achieved at a given load for higher temperatures, due to thermally activated creep mechanisms. The observation that the curves of tensile stress versus minimum creep rate are nearly parallel at the three temperatures is indicative of a common apparent activation energy and underlying creep mechanism, or set of creep mechanisms, that controls deformation at each of these temperatures and strain rates.

During the strain hardening process, the dislocation density gradually increases, leading to a decreasing creep strain rate as dislocation mobility is impeded. Eventually, during secondary creep, a steady-state is achieved between the competing processes of recovery (dislocation annihilation) and

strain hardening (dislocation generation), where the ratio of the rates of these two competing processes yields the minimum creep strain rate.²⁴ However, the treatment of strain hardening and, consequently, transient creep strain is not consistent in the literature. The saturated transient creep strain has been assumed to be a constant over all temperatures for Sn3.8Ag0.7Cu. It has also been considered to be solely a function of minimum strain rate for Sn4.0Ag0.5Cu.²⁵ In both cases the strain hardening was assumed not to be thermally activated. In contrast to these two assumptions, Fig. 10 shows that transient saturation creep strain, $\epsilon_{tr,sat}$, for Sn3.8Ag0.7Cu was a function of both the minimum creep strain rate and temperature. Larger values of saturated transient creep strain were measured at lower temperatures, while there was a gradual increase in saturated transient creep strain at higher minimum creep strain rates for a given temperature.

Strain hardening is commonly thought to result from dislocation pile-up at barriers (planar glide), occurring over a relatively large number of interatomic distances and, consequently, being a function of strain rate but independent of temperature.²⁶ However, strain hardening may also occur through the intersection of dislocations, termed dislocation forests (wavy glide), where dislocations moving in the slip plane cut through other dislocations intersecting the active slip plane.^{4,26} This hardening mechanism is enhanced by cross-slip that accelerates the dislocation rearrangement process in high stacking fault energy materials. Strain hardening by this cutting process arises from short-range forces occurring over distances less than 5 to 10 interatomic distances, and, therefore, it is subject to thermal recovery and is both strain rate and temperature dependent. In contrast, strain hardening arising from dislocation pile-up at barriers occurs over much longer interatomic distances and is, therefore, relatively independent of strain rate and temperature. Therefore, the results of Fig. 10 suggest that strain hardening through dislocation climb and the cutting of intersecting dislocation forests is active in Sn3.8Ag0.7Cu. Generally, metals with high stacking fault energy, such as tin, show only a very small propensity to harden by dislocation pile-up, because they can deform easily by cross-slip. Since Sn-rich lead-free solders are principally composed of a β -Sn matrix, they exhibit strain hardening deformation characteristic of high stacking fault energy pure metals.²³

Stress Relaxation

Knowledge of the elastic modulus of the tensile specimen and the machine compliance of the load-bearing frame were essential to the correct interpretation of the stress relaxation test. Young's modulus can vary widely, depending on the rate of loading and the scale of deformation. For example, the

Table II. Stress Relaxation Experimental Parameters

Temperature, T (°C)	25	75	100
Crosshead speed, \dot{X} (cm/min)	2	4	6
Static Young's modulus, E (GPa)	46	30	19
Specimen-machine elastic modulus, C (GPa)	1.33	3.13	12.43

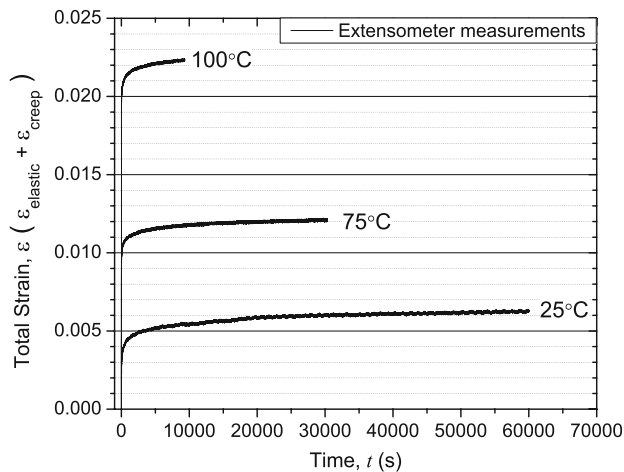


Fig. 11. Extensometer measured total strain versus time (25°C, 75°C, 100°C). Fast crosshead speed 2.0/4.0/6.0 cm/min at 25/75/100°C.

dynamic modulus measured by the propagation of an acoustic wave through the material is associated with purely elastic deformation²⁷ and may be significantly greater than the static modulus measured from the stress-strain response of the solder loaded in tension at the start of the stress relaxation test. In both cases, Young's modulus decreases with increasing temperature, because, at higher temperatures, atoms in the lattice move farther apart due to thermal expansion lowering interatomic bonding forces.²⁶ The static modulus values measured at the start of the stress relaxation tests (Table II) are similar to those reported in Ref. 28 for Sn3.8Ag0.7Cu bulk solder tensile specimens.

The combined specimen-machine elastic modulus was calculated with Eqs. 5–8, where total strain rate and load rate inputs were also calculated. The total strain rate ($\dot{\epsilon}_{\text{elastic}} + \dot{\epsilon}_{\text{plastic}}$) was computed from the total strain versus time response in Fig. 11. Similarly, the load rate (\dot{P}) was computed from the force versus time response in Fig. 12. The noisy experimental data for total strain and load responses versus time were each fit with a three-parameter logarithmic expression given by Eq. 9,²⁹ where y is either total strain or load; X is time, and α , β , and γ are fitting parameters. Accordingly, the creep strain rate was subsequently calculated by subtraction of the elastic component from the total strain rate.

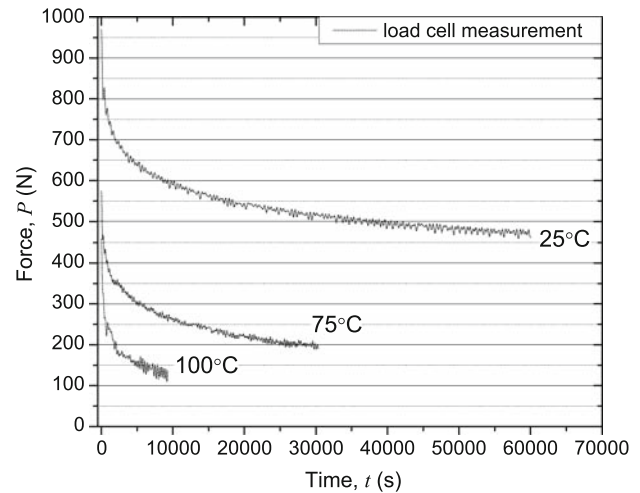


Fig. 12. Force versus time from stress relaxation tests (25°C, 75°C, 100°C). Fast crosshead speed 2.0/4.0/6.0 cm/min at 25/75/100°C.

$$y = \alpha - \beta * \log(X + \gamma) \quad (9)$$

As shown in Fig. 11, the total strain in the gauge section was not constant during these fixed displacement stress relaxation tests, but rather it increased with time over all test temperatures. Owing to faster creep rates at higher temperatures, it took significantly less time for the stress relaxation to approach the practical resolution limit of the test ($1 \times 10^{-8} \text{ s}^{-1}$), hence the discrepancy in test duration at 25°C, 75°C, and 100°C. As anticipated, more strain was measured at higher temperatures as a result of more elastic strain during crosshead movement and higher creep rates. This is also evident in Fig. 12, where the larger strains at higher temperatures coincide with periods of faster relaxation.

Figure 13 shows the creep strain rate versus stress from the stress relaxation tests together with the minimum creep strain rate versus stress from the constant load creep tests. It is evident that the creep strain rate for stress relaxation was faster than the minimum creep strain rate for the corresponding stress measured from constant load creep tests. This was simply a result of the steady-state and transient creep both contributing to deformation during stress relaxation. The magnitude of the strain rate offset between stress relaxation and constant load creep results is indicative of the relative importance of transient creep, and Fig. 13 shows that the transient contributions were not uniform over all temperature and strain rates; i.e., the strain rate difference increased as the temperature fell and, to a lesser extent, as the stress decreased. This is consistent with Fig. 10, which shows that the largest saturation transient creep strain among all constant load creep tests occurred at 25°C, while the smallest values were measured at 100°C. For all three test temperatures the amount of saturation transient creep strain increased with increasing minimum creep rate.

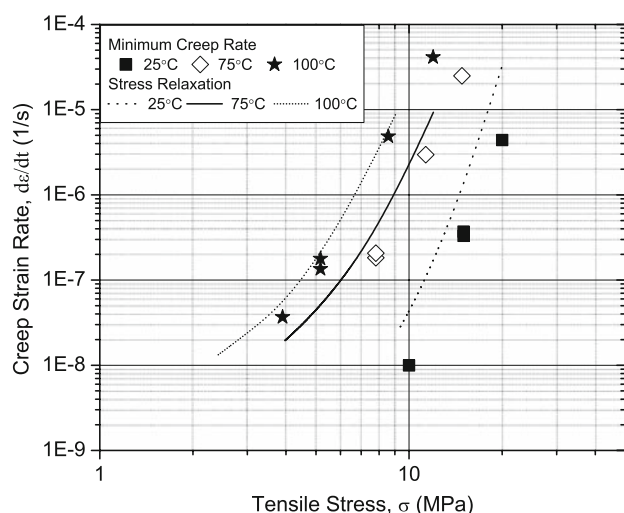


Fig. 13. Creep strain rate versus tensile stress from creep and stress relaxation tests (25°C, 75°C, 100°C).

During stress relaxation, stress decreases over time. During the finite moments between incremental load drops, there may be insufficient time to approach transient creep saturation; therefore, the strain rate is higher than it would be under purely steady-state creep. The extent to which transient creep saturation is approached during the stress relaxation test changes as a function of temperature and strain rate. In Fig. 13, for 25°C, the entire stress relaxation response has a significant transient component and does not appear to approach saturation at any point. However, at 75°C, at the highest strain rates ($1 \times 10^{-5} \text{ s}^{-1}$), the strain rate offset between the stress relaxation and constant load curves is relatively small, and, at incremental load drops, transient creep saturation may be approached. This changes for the same temperature at lower strain rates ($1 \times 10^{-7} \text{ s}^{-1}$), where transient saturation is now much further away at each incremental load drop, hence the larger strain rate offset. If we consider 100°C curves, the strain rate offsets are indistinguishable at the highest strain

rates ($1 \times 10^{-5} \text{ s}^{-1}$) through to nearly ($1 \times 10^{-7} \text{ s}^{-1}$), and only at strain rates below this point does the strain rate offset become notable. This implies that, during a large portion of the 100°C stress relaxation response, saturation transient creep strain was approached at each incremental load drop.

These observations suggest that the proportion of transient and steady-state creep changes as a function of temperature and strain rate. Thus, a constitutive model that considers only steady-state creep when fit to a stress relaxation response for SnAgCu will have some unknown and variable error when applied to temperatures and strain rates beyond its original data set, because it does not account for the effect of the changing proportions of transient and steady-state contributions. Such a finding was reported in Ref. 30 where a steady state creep model was fit to creep and stress relaxation data for Sn3.5Ag eutectic solder. Different sets of parameters, including stress exponent and activation energy, were required to fit the creep and stress relaxation measurements. The discrepancy was attributed to a complex dependency on stress and strain rate beyond the capability of the steady state creep model.

Microstructural Development

The microstructure shown in Fig. 14 resulted from solder that was cast and cooled at 2°C/s prior to aging at $\frac{2}{3}T_m$ (54°C) for 16 h—the same thermal history used for all creep and stress relaxation specimens. Figure 3 shows that the pre-test aging period purged the β -Sn matrix of most fine dispersions of Ag₃Sn and Cu₆Sn₅ as the intermetallics coarsened forming rod-like particle groupings and binary and eutectic solid solution regions. After isothermal creep testing at 100°C and an applied engineering stress of 8.6 MPa, the microstructure became more complex (Fig. 14). From the corresponding creep curve (Fig. 8), sufficient time had elapsed for the microstructure to reflect a saturation of transient creep and a complete transition to a steady-state creep rate.

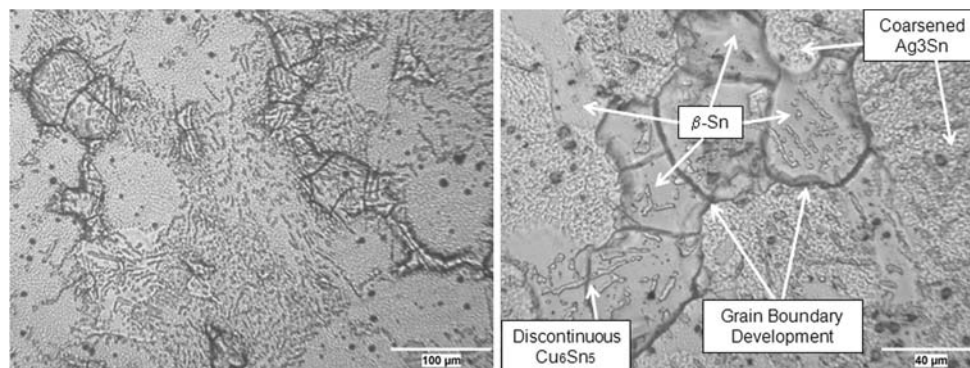


Fig. 14. Optical micrographs of multiple locations using quarter-wave plate of aged and crept specimen microstructure, $\sigma = 8.6 \text{ MPa}$, $T = 100^\circ\text{C}$. (a) Low magnification, (b) high magnification.

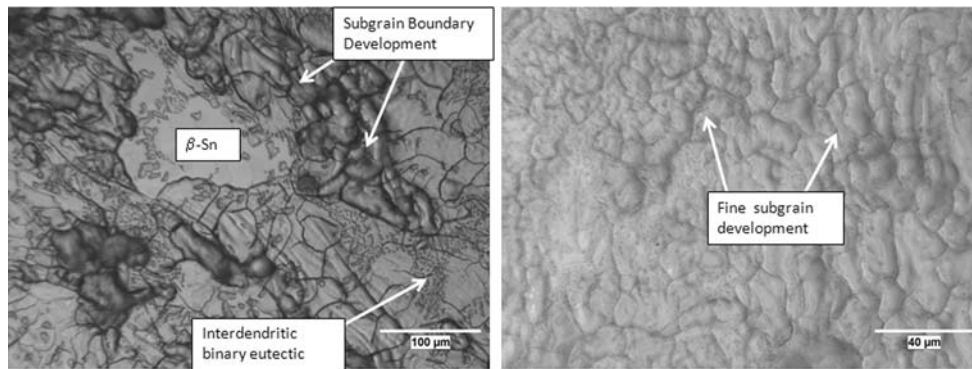


Fig. 15. Optical micrographs of multiple locations using quarter-wave plate of aged and crept specimen microstructure, $\sigma = 20$ MPa, $T = 25^\circ\text{C}$. (a) Low magnification, (b) high magnification.

One notable difference between the microstructure of specimens before and after creep is the appearance of substructure development in individual grains. During primary creep, intragranular creep deformation of the β -Sn matrix by the dislocation climb mechanism causes a rearrangement of dislocations into a lower-energy configuration (less local lattice strain) of low-angle boundaries consisting of dislocation arrays that ultimately form the polygon-like network of subgrains seen in Fig. 15. This polygon-like network formation process is prominent in high stacking fault energy materials such as Sn and Sn-rich lead free solders when the temperature and extent of deformation are low enough to prevent the spontaneous formation of new grains by recrystallization.^{4,31} Subgrain formation and recrystallization are two competitive processes where the occurrence of the former reduces the tendency of the latter.³² Recrystallization is not evident in Fig. 15 and probably had minimal impact during the creep and stress relaxation tests, given the low creep strain deformation measured.

In Fig. 14 the absence of void coalescence at boundaries, and the lack of trans-granular or intra-granular cracking, is an indication that grain boundary sliding did not contribute significantly to creep deformation in these experiments. Furthermore, the subgrain contains β -Sn with a distribution of Cu₆Sn₅ particles that also cross the low-angle boundaries. In a few instances the Cu₆Sn₅ particles appeared discontinuous at the boundary, but each portion of the particle remained aligned across the boundary, indicating that grain boundary sliding was absent. More commonly, the Cu₆Sn₅ particles that spanned the low-angle boundaries remained clearly continuous across the boundary. The presence of primary phase Cu₆Sn₅ and Ag₃Sn particles serves to resist defect movement, especially when located along the grain boundary.³ However, the effect of Ag₃Sn and Cu₆Sn₅ primary phases in improving creep resistance is limited for diffusion-based mechanisms because of the contribution of thermally activated vacancy diffusion to mass transport, whether alone or in assisting dislocation

climb. In general, diffusion-based creep mechanisms (i.e., self-diffusion and core diffusion), either alone or in combination, were expected to dominate in our experiments because of the high homologous temperature ($T > 0.6T_m$).

Figure 15 shows the microstructure of a specimen tested under constant load creep at 25°C with an applied engineering stress of 20 MPa. In contrast to the creep test at 100°C (Fig. 14), where the total creep strain was less than 2%, the microstructure of Fig. 15 corresponds to a total creep strain of approximately 6% at room temperature. The lower temperature resulted in a higher dislocation density, with more subgrains being created during primary creep. At room temperature, additional strain hardening occurred during primary creep because the thermally activated recovery processes played a smaller role. This substructure development is the reason a larger saturation transient creep strain was measured at lower temperatures and at higher strain rates.

Strain hardening is caused by the penetration of dislocations into subgrain boundaries that increase the local dislocation density and form a barrier to dislocation motion. The competing recovery process is caused by a lowering of the dislocation density through mutual climb and annihilation of edge dislocations.³³ The dislocation density increases directly as a function of increasing stress, and the creep rate is proportional to the dislocation density.²³ This effect was observed in constant load creep tests where faster minimum creep rates were measured with higher dislocation densities at higher stresses. At the start of the stress relaxation test a high stress was obtained, creating greater dislocation density than would be obtained for a constant load creep test at lower stress levels. This elevated initial dislocation density accounted for the faster measured creep rates at a given stress in the stress relaxation tests than the corresponding minimum creep rate from the constant load tests. The smaller relative offset in creep rates between these two tests at higher temperatures can be accounted for by the increased thermal recovery that decreases the

dislocation density and reduces the creep strain rate during stress relaxation.

The role of thermal activation in the polygon-like network formation that underlies strain hardening during primary creep in SnAgCu has been included in a constitutive creep model presented in a companion paper.¹⁰

CONCLUSIONS

Both transient and steady-state creep play a significant role in creep deformation over a broad range of strain rates ($1 \times 10^{-8} < \dot{\epsilon}_{ss} < 1 \times 10^{-4}$) and temperatures (25°C, 75°C, and 100°C). Matrix creep and dislocation core diffusion, either alone or in combination, dominate these strain rates and temperatures.

The extent of strain hardening during primary creep and, hence, the magnitude of the saturation transient creep strain was a function of temperature and strain rate.

The measured strain rates from stress relaxation were consistently higher than the minimum creep rates measured during constant-load creep testing, the difference being a strong function of temperature and, to a lesser extent, strain rate.

Transient creep strain can be a significant portion of total creep strain in Sn-rich lead-free solders such as Sn3.8Ag0.7Cu. The relative amount of transient creep is strongly dependent on the temperature and strain rate. Cyclic loading applications, like accelerated thermal cycling, are affected by primary creep because relatively rapid temperature and/or stress changes occur throughout a given cycle.

ACKNOWLEDGEMENTS

The authors thank Research in Motion Ltd. (Laura Turbini, Bev Christian, Gene Burger), Celestica Inc. (Polina Snugovsky, Peter Arrowsmith, Rob Emery), University of Toronto, Ontario Centres of Excellence (OCE), and National Sciences and Engineering Research Council of Canada (NSERC) for their invaluable support of this research.

REFERENCES

1. W.W. Lee, L.T. Nguyen, and G.S. Selvaduray, *Microelectron. Reliab.* 40, 231 (2000).
2. D.R. Shirley, H.R. Ghorbani, and J.K. Spelt, *Microelectron. Reliab.* 48, 455 (2008).
3. P. Vianco, *Lead-Free Solder Interconnect Reliability*, ed. D. Shangguan (Materials Park, OH: ASM, 2005), p. 67.
4. G.E. Dieter, *Mechanical Metallurgy* (New York: McGraw-Hill, 1986).
5. A.J. Kennedy, *Processes of Creep and Fatigue in Metals* (New York: Wiley, 1963).
6. L.J. Ladani, *IEEE Trans. Device Mater. Res.* 8, 375 (2008).
7. X. Nie, D. Bhate, D. Chan, W. Chen, G. Subbarayan, and I. Dutta, *11th Intersociety Conference on IEEE ITherm* (2008), p. 676.
8. Q. Zhang, A. Dasgupta, D. Nelson, and H. Pallavicini, *J. Electron. Packag.* 127, 415 (2005).
9. G. Cuddalorepatta and A. Dasgupta, *IEEE EuroSimE* (Germany, 2008), p. 1.
10. D.R. Shirley and J.K. Spelt, *J. Electron. Mater.* (2009). doi: [10.1007/s11664-009-0906-0](https://doi.org/10.1007/s11664-009-0906-0).
11. ISO 204:1997 Uninterrupted Uniaxial Creep Testing (1997).
12. ASTM E139-00 Standard Test Methods for Conducting Creep (2000), p. 285.
13. L. Ma, S. Bao, D. Lv, Z. Du, and S. Li, *IEEE International Conferences on Electron Packaging Technology* (Shanghai, 2007), p. 1.
14. W.J. Plumbridge, R.J. Matela, and A. Westwater, *Structural Integrity and Reliability in Electronics* (Boston: Kluwer, 2003).
15. A.D. Holmes and F. Jing, *J. Phys. D: Appl. Phys.* 27, 2475 (1994).
16. I. Koychev and R. Lothian, *Research and Development in Intelligent Systems XXII Proceedings of AI-2005* (Cambridge, UK, 2005), p. 46.
17. M. Kukar, *Artificial Intelligence in Medicine in Europe* (Protaras, Cyprus, 2003), p. 355.
18. E.O. Brigham, *The Fast Fourier Transform and Its Applications* (Englewood Cliffs, NJ: Prentice Hall, 1988).
19. D.F. Elliott and K.R. Rao, *Fast Transforms: Algorithms, Analyses, Applications* (New York: Academic Press, 1982).
20. C. Pan, *IEEE Trans. Signal Process.* 49, 444 (2001).
21. J.E. Hockett and P.P. Gillis, *Int. J. Mech. Sci.* 13, 251 (1971).
22. J.H. Holbrook, J.C. Swearengen, and R.W. Rohde, *ASTM STP* 765, 80 (1982).
23. O.D. Sherby and P.M. Burke, *Prog. Mater. Sci.* 13, 323 (1968).
24. S.O. Ojedian and O. Ajaja, *J. Mater. Sci.* 23, 4037 (1988).
25. A. Schubert, R. Dudek, E. Auerswald, A. Gollhardt, B. Michel, and H. Reichl, *IEEE ECTC* (New Orleans, 2003), p. 603.
26. R.W. Hertzberg, *Deformation and Fracture Mechanics of Engineering Materials* (New York: Wiley, 1996).
27. Y. Nakamura and T. Ono, *J. Mater. Sci.* 40, 3267 (2005).
28. J.H.L. Pang, B.S. Xiong, C.C. Neo, X.R. Zhang, and T.H. Low, *ECTC* (New Orleans, 2003).
29. D.A. Ratkowsky, *Handbook of Nonlinear Regression Models* (New York: Marcel Dekker, 1990).
30. H. Mavoori, J. Chin, S. Vaynman, B. Moran, L. Keer, and M. Fine, *J. Electron. Mater.* 26, 783 (1997).
31. P. Gay and A. Kelly, *Acta Crystallogr.* 6, 172 (1953).
32. D. Hardwick, C.M. Sellars, and W.J. Tegart, *J. Inst. Met.* 90, 21 (1961).
33. W.R. Thorpe and I.O. Smith, *Phys. Status Solidi* 52, 487 (1979).